Complete genome sequence of an African swine fever virus (ASFV POL/2015/Podlaskie) determined directly from pig erythrocyte-associated nucleic acid

Olesen, Ann Sofie; Lohse, Louise; Dalgaard, Marlene Danner; Woniakowski, Grzegorz; Belsham, Graham J.; Bøtner, Anette; Rasmussen, Thomas Bruun

Published in:
Journal of Virological Methods

Link to article, DOI:
10.1016/j.jviromet.2018.07.015

Publication date:
2018

Document Version
Peer reviewed version

Title: Complete genome sequence of an African swine fever virus (ASFV POL/2015/Podlaskie) determined directly from pig erythrocyte-associated nucleic acid

Authors: Ann Sofie Olesen, Louise Lohse, Marlene Danner Dalgaard, Grzegorz Woźniakowski, Graham J. Belsham, Anette Bøtner, Thomas Bruun Rasmussen

PII: S0166-0934(18)30195-2
DOI: https://doi.org/10.1016/j.jviromet.2018.07.015
Reference: VIRMET 13509

To appear in: *Journal of Virological Methods*

Received date: 9-4-2018
Revised date: 27-7-2018
Accepted date: 27-7-2018

Please cite this article as: Olesen AS, Lohse L, Dalgaard MD, Woźniakowski G, Belsham GJ, Bøtner A, Rasmussen TB, Complete genome sequence of an African swine fever virus (ASFV POL/2015/Podlaskie) determined directly from pig erythrocyte-associated nucleic acid, *Journal of Virological Methods* (2018), https://doi.org/10.1016/j.jviromet.2018.07.015

Complete genome sequence of an African swine fever virus (ASFV POL/2015/Podlaskie) determined directly from pig erythrocyte-associated nucleic acid

Ann Sofie Olesen [a], Louise Lohse [a], Marlene Danner Dalgaard [b], Grzegorz Woźniakowski [c], Graham J. Belsham [a], Anette Bøtner [a], Thomas Bruun Rasmussen [a*]

[a]*DTU National Veterinary Institute, Technical University of Denmark, Lindholm, Kalvehave, Denmark*

[b]*DTU Bioinformatics, Technical University of Denmark, Kgs. Lyngby, Denmark*

[c]*Department of Swine Diseases, National Veterinary Research Institute, Pulawy, Poland*

[*] **Corresponding author**: Thomas Bruun Rasmussen, DTU National Veterinary Institute, Technical University of Denmark, Lindholm, DK-4771 Kalvehave, Denmark, phone number +45 35887850, e-mail address tbrur@vet.dtu.dk

Highlights:

- The complete genome sequence of an ASFV from Poland (2015) was determined.
- This sequence was obtained directly from pig erythrocyte-associated nucleic acids.
- The work should assist studies on the genetic diversity of the European ASFVs.

ABSTRACT

African swine fever (ASF) is an important disease of domestic pigs and wild boar. The disease is caused by African swine fever virus (ASFV). In 2014, ASFV was introduced into Eastern Europe, and it has since then continued to spread within various Eastern European countries. Investigating differences in sequences between ASFV isolates may be a valuable tool to understand differences in virulence among them, however currently, no complete genome sequences of the viruses responsible for the Eastern European outbreaks have been reported. In this study, the complete genome sequence of a highly virulent ASFV was determined directly from erythrocyte-associated nucleic acids obtained from a pig experimentally infected with an isolate from Poland (ASFV POL/2015/Podlaskie). The sequence (ca. 189kb) of this recent European ASFV showed 95 nt differences (99.95 % identity) from the ASFV Georgia 2007/1 genome. The complete sequence of ASFV/Pol/2015/Podlaskie should assist further studies on the genetic diversity and evolution of the European ASFVs.

Keywords: ASF, complete genome, NGS, Poland, WGS

African swine fever (ASF) is an important viral disease of domestic pigs and wild boar (EFSA Panel on Animal Health and Welfare, 2014). The African swine fever virus (ASFV), the only member of the *Asfarviridae* family, has a large (170-193 kbp) double-stranded DNA genome (Dixon et al., 2013). In 2007, ASFV was introduced into Georgia, and it has then spread within the Transcaucasian countries, the Russian Federation, Belarus and Ukraine (EFSA Panel on Animal Health and Welfare, 2014). More recently, since 2014, the virus has spread into various Eastern European countries (EFSA Panel on Animal Health and Welfare, 2015). Currently, virulence factors of the virus are not fully defined and

filling these knowledge gaps is required for understanding differences in virulence among ASFV strains (Blome et al., 2013). Comparison of genome sequences between strains of different virulence may be a valuable tool to understand such issues. At present, only the complete genome sequence of the ASFV Georgia 2007/1 isolate is available (Chapman et al., 2011). No complete genome sequences of viruses responsible for the Eastern European ASF outbreaks have been reported. In this study, we describe the complete genome sequence of the recent Eastern European ASFV POL/2015/Podlaskie (the isolate is described in Olesen et al. (2017)).

The ASFV POL/2015/Podlaskie sequence was determined using blood from a domestic pig experimentally infected with this virus isolate; this was derived from spleen material collected from an ASFV-infected wild boar in Podlaskie voivodeship (province) in February 2015 (Olesen et al., 2017). For the experimental infection, virus was isolated from spleen material by two passages in porcine pulmonary alveolar macrophages. Following virus inoculation, EDTA-stabilized blood was obtained from the pig (no. 23) at euthanasia, eight days later, following an acute rapid disease progression characterized by fever, incoordination, diarrhea and viremia (Olesen et al., 2017). Within 12 hours of collection of the blood sample, erythrocytes were separated from mononuclear cells and the plasma fraction of the blood using LymphoPrep™ density gradient medium (STEMCELL™ Technologies). Viral DNA was extracted from the erythrocyte-enriched sample (9.5 $\log_{10}$ virus genome copies/mL sample) on a MagNA Pure 96 system (Olesen et al., 2017). DNA libraries were generated using the extracted DNA with the Nextera XT DNA library preparation kit and sequenced on a MiSeq with Reagent Kit v3 600 bp (Illumina Inc.). The complete ASFV genome was determined by *de novo* assembly and iterative mapping using CLC Genomics Workbench (QIAGEN). PCR products were prepared and sequenced using Sanger and MiSeq methodologies to confirm parts of the sequence representing ca. 35kb of the genome. The PCR products were amplified from DNA extracted from the blood of pig 23, but also

from spleen material collected from a pig experimentally infected with a 2007 Georgian isolate (Nielsen and Uttenthal, 2013). Finally, the consensus sequence was compared to the ASFV Georgia 2007/1 genome (accession number FR682468 (Chapman et al., 2011)) using CLC Genomics Workbench and Geneious version 10.2.3 (Biomatters).

The complete genome sequence of ASFV POL/2015/Podlaskie (189393 bp, GenBank accession number MH681419) was obtained from a reference-based alignment consisting of 113246 mapped reads (0.84 % of 13489946 total reads) with an average coverage of 103 reads per nt using blood from pig 23. In addition, an identical consensus sequence, with a coverage of 159 reads per nt, was assembled from 253045 mapped reads combined from this pig and ten other pigs infected with the same inoculum (as described in Olesen et al. (2017)) using the same approach.

Comparison to the ASFV Georgia 2007/1 sequence revealed just 95 nt differences (thus 99.95 % identity) (Table 1). Of note, were a 3 bp deletion leading to loss of the terminal glycine codon in the MFG_110-14L open reading frame (ORF), and a deletion of a stretch of 6 Gs within the ORF of the ASFV_G_ACD-00350 gene resulting in the loss of two glycine codons. The effect on protein functions resulting from these deletions is unknown. Furthermore, four non-synonymous mutations were observed at: nt 62197 in the gene F1055L (Glu to Val), nt 155895 in the H240R gene (Arg to His), nt 166066 in the gene E199L (Gly to Arg), and nt 181965 in the gene I9R (Lys to Glu). Eleven other SNPs were observed, these were synonymous or situated in non-coding regions, including heterogeneity at nt 189223 (G/A), nt 189226 (A/G) and nt 189253 (C/G) (Table 1).

The remaining differences were 1-2 nt indels in homopolymeric regions, predominantly within intergenic regions, but 17 of these indels were identified within coding regions (Table 1). Some indels (17) were confirmed by sequencing of PCR products to rule out possible homopolymer artefacts. Most of these indels (14 out of 17) were also observed in DNA extracted from the spleen material collected from the

pig experimentally infected with a 2007 Georgian isolate (Nielsen and Uttenthal, 2013) (Table 1). This indicates that different variants of the ASFV Georgia 2007/1 sequence exist; these should be investigated further.

The complete genome sequence of the ASFV/Pol/2015/Podlaskie obtained from an experimentally infected pig should assist further studies on the genetic diversity and the evolution of the European ASFVs. Importantly, the use of whole-genome sequencing of viruses directly from blood-derived nucleic acid samples should minimize bias introduced by steps such as virus isolation and PCR amplification prior to sequencing.

Conflict of interest

The authors declare no conflicts of interest.

Acknowledgements

We are grateful to laboratory technicians Fie Fisker Brønnum, Hanne Egelund Hansen, Hanne Petersen, Heidi Elbrink and Neslihan Bicen, and to animal caretakers Heidi Lehman, Janni Oxfeldt, Josefine Cecilie Hammer and Marion Baltzer Petersen for their excellent work.

We gratefully acknowledge financial support from the CoVetLab, the Danish Ministry of Environment and Food and the Technical University of Denmark. The National Science Centre supported the scientific activities of Dr. Grzegorz Woźniakowski (grant. no*:* UMO-2016/21/D/NZ6/00974).

References

Blome, S., Gabriel, C., Beer, M., 2013. Pathogenesis of African swine fever in domestic pigs and European wild boar. Virus Res. 173, 122-130.

Chapman, D.A.G., Darby, A.C., Da Silva, M., Upton, C., Radford, A.D., Dixon, L.K., 2011. Genomic analysis of highly virulent isolate of African swine fever virus. Emerg. Infect. Dis. 17, 599-605.

Dixon, L.K., Chapman, D.A.G., Netherton, C.L., Upton, C., 2013. African swine fever virus replication and genomics. Virus Res. 173, 3-14.

EFSA Panel on Animal Health and Welfare, 2014. Scientific Opinion on African swine fever. EFSA J. 12 (3628), 1-77.

EFSA Panel on Animal Health and Welfare, 2015. Scientific Opinion on African swine fever. EFSA J. 13 (4163), 1-92.

Nielsen, J., Uttenthal, Å., 2013. Hvordan ser afrikansk svinepest ud i danske grise? Dansk Veterinærtidsskrift 12, 20-22.

Olesen, A.S., Lohse, L., Boklund, A., Halasa, T., Gallardo, C., Pejsak, Z., Belsham, G.J., Rasmussen, T.B., Bøtner, A., 2017. Transmission of African swine fever virus from infected pigs by direct contact and aerosol routes. Vet. Microbiol. 211, 92-102.

Table

Table 1

Variation detected between ASFV POL/2015/Podlaskie and ASFV Georgia 2007/1.

Type	Genome position(*)(**) and modification	Annotation (ORF position(*))	Effect on ORF(***)	PCR (POL/2015)	PCR (Geo/07)
Indel	412_413insA	Terminal inverted repeat (1..420)		insA	insA
Indel	432_433insCC			insCC	No change
Indel	433_434_insT			insT	InsT
Indel	440_441insA			InsA	InsA
Indel	1604_1607delTT	MGF_360-1L (852..1934)	Frameshift		
Indel	1622_1623insT	MGF_360-1L (852..1934)	Frameshift		
Indel	3246_3247insA	KP177R (3212..3781)	Frameshift		
Indel	3861_3862insA				
Indel	5670_5671_insA	ASFV_G_ACD_00070 (<5604..5732)	Frameshift		
Indel	5820_5821insT				
Indel	5869_5870insA				
Indel	7823_7824insA	ASFV_G_ACD_00120 (7795..7920)	Frameshift		
Indel	7940_7941_insA	ASFV_G_ACD_00120 (7795..7920)	Frameshift		
Indel	8505_8506insA				
Indel	10465_10466insA				

Indel	13266_13270delCCC	MGF_110-14L (13253..13603)	Glycine codon deleted(****)		
Indel	13665_13666insA	ASFV_G_ACD_00240 (13602..13685)	Frameshift		
Indel	16879_16880insGG				
Indel	18736_18737insT				
Indel	18830_18831insGG				
Indel	19032_19039delGGGGGG	ASFV_G_ACD_00350 (18994..19128)	Two Glycine codons deleted	delGGGGGG	No change
Indel	19461_19462insT			InsT	InsT
Indel	20669_20670insA				
Indel	20834_20836delG				
Indel	21588_21589insA				
Indel	23027_27028insA			InsA	InsA
SNP	A26037G	MGF_360-10L (25401..26438)	Synonymous mutation	G	No change
Indel	26455_26456insT			InsT	InsT
Indel	31678_31679insA			InsA	InsA
Indel	39024_39025insTT			InsTT	InsT
Indel	40730_40731insT			InsT	InsT
Indel	42616_42617insT			InsT	InsT
Indel	44547_44548insT			InsT	InsT
Indel	47321_47322insA			InsA	InsA
Indel	48663_48664insA	A151R (48651..49127)	Frameshift	InsA	InsA
Indel	55974_55975insA				
SNP	T62197A	F1055L (59853..62999)	Non synonymous mutation (E to V)		

Indel	63845_63846insT				
Indel	63860_63861insA				
Indel	63867_63868insT				
Indel	63876_63877insT				
SNP	G68524A	EP1242L (66514..70242)	Synonymous mutation		
Indel	81525_81526insT	C84L (81353..81592)	Frameshift		
Indel	81619_81620insA				
SNP	G99635A	B34L (99614..100678)	Synonymous mutation	A	No change
Indel	113166_113167insA				
Indel	117181_117182insT				
Indel	124704_124705insA				
Indel	124803_124804insTT	CP204L (124770..125375)	Frameshift		
Indel	133463_133464insA				
Indel	138421_138422insT	D129L (138420..138809)	Synonymous mutation		
Indel	143757_143758insA				
Indel	152792_152793insA				
SNP	C155227T				
SNP	G155895A	H240R (155339..156064)	Non synonymous mutation (R to H)		
Indel	161251_161253delA	QP383R (160525..161676)	Frameshift (change of 12 contiguous residues)		

Indel	161292_161293insA	QP383R (160525..161676)	Frameshift (return to initial ORF)		
SNP	C165718T				
SNP	C165762T				
SNP	C165780T				
SNP	C165782T				
SNP	C166066T	E199L (165845..166444)	Non synonymous mutation (G to R)		
Indel	167763_167764insA				
Indel	169454_169455insT				
Indel	169462_169463insA				
Indel	169741_169742insT	I267L (169605..170444)	Frameshift		
Indel	172017_172018insT				
Indel	173565_173566insA				
Indel	174954_174955insA	ASFV_G_ACD_01760 (174920..175003)	Frameshift (fusion with downstream ORF I177L)		
Indel	175632_175633insA				
Indel	177493_177494insA	MGF_360-16R (176590..177519)	Frameshift (fusion with downstream ORF DP63R)		
Indel	179379_179380insA				
SNP	A181965G	I9R (181698..181988)	Non synonymous mutation (K to E) near C-terminus		

Indel	183305_183306insT				
Indel	183314_183315insT				
Indel	184967_184969delT	ASFV_G_ACD_01940 (184949..185080)	Frameshift		
Indel	186352_186353insT				
Indel	189022_189024delA	DP60R (188997..189161)	Frameshift	delA	No change
SNP	G189223A		Quasispecies	G/A	G/A
SNP	A189226G		Quasispecies	No change	No change
SNP	C189253G		Quasispecies	C/G	C/G
Indel	189282_189283insT			InsT	InsT

(*) The positions of detected variations and open reading frame (ORF) positions are based on the reference genome (ASFV Georgia 2007/1, FR682468).

(**) The position of the indels refer to the position of the last nucleotide within each homopolymer region. Genome positions for indels are given by the flanking nt positions.

(***) The effect of the frameshifts can increase or decrease the size of the protein or can just alter the amino acid sequence.

(****) Read in right to left direction.